

1611  
#G/10W  
04-09-02  
RECEIVED

FEB 20 2002

OIPE

TECH CENTER 1600/2900

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/963,521

DATE: 12/20/2001

TIME: 09:24:45

Input Set : A:\23282413.app

Output Set: N:\CRF3\12192001\I963521.raw

ENTERED

```

3 <110> APPLICANT: ZIEGLER, PETRA
4   EGDELING, LOTHAR
5   SAHM, HERMANN
7 <120> TITLE OF INVENTION: NEW NUCLEOTIDE SEQUENCES CODING FOR THE THRE GENE
8   AND PROCESS FOR THE ENZYMATIC PRODUCTION OF
9   L-THREONINE USING CORYNEFORM BACTERIA
11 <130> FILE REFERENCE: 21123/282413/MAS
13 <140> CURRENT APPLICATION NUMBER: 09/963,521
14 <141> CURRENT FILING DATE: 2001-09-27
16 <150> PRIOR APPLICATION NUMBER: 09/431,099
17 <151> PRIOR FILING DATE: 1999-11-01
19 <150> PRIOR APPLICATION NUMBER: DE 199 41 478.5
20 <151> PRIOR FILING DATE: 1999-09-01
22 <160> NUMBER OF SEQ ID NOS: 10
24 <170> SOFTWARE: PatentIn Ver. 2.1
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 2817
28 <212> TYPE: DNA
29 <213> ORGANISM: Corynebacterium glutamicum
31 <220> FEATURE:
32 <221> NAME/KEY: CDS
33 <222> LOCATION: (398)..(1864)
34 <223> OTHER INFORMATION: thrE-Gen
36 <400> SEQUENCE: 1
37 aatgaaataa tccccacacc aactggcgac attcaaacac cgtttcattt ccaaaccatcg 60
39 agccaaggga aaagaaagcc cctaagcccc gtgttattaa atggagactc tttggagacc 120
41 tcaagccaaa aaggggcatt ttcattaaga aaataccctt ttgacctggt gttattgagc 180
43 tggagaagag acttgaactc tcaacctacg cattacaagt gcgttgcgct gccaatgagc 240
45 ccactccagc accgcagatg ctgatgatca acaactacga atacgtatct tagcgtatgt 300
47 gtacatcaca atggaattcg gggctagagt atctggtgaa ccgtgcataa acgacctgtg 360
49 attggactct ttttccttgc aaaatgtttt ccagcgg atg ttg agt ttt gcg acc 415
50                                     Met Leu Ser Phe Ala Thr
51                                     1           5
53 ctt cgt ggc cgc att tca aca gtt gac gct gca aaa gcc gca cct ccg 463
54 Leu Arg Gly Arg Ile Ser Thr Val Asp Ala Ala Lys Ala Ala Pro Pro
55          10          15          20
57 cca tcg cca cta gcc ccg att gat ctc act gac cat agt caa gtg gcc 511
58 Pro Ser Pro Leu Ala Pro Ile Asp Leu Thr Asp His Ser Gln Val Ala
59          25          30          35
61 ggt gtg atg aat ttg gct gcg aga att ggc gat att ttg ctt tct tca 559
62 Gly Val Met Asn Leu Ala Ala Arg Ile Gly Asp Ile Leu Leu Ser Ser
63          40          45          50
65 ggt acg tca aac agt gat acc aag gtg caa gtt cga gcg gtg acc tct 607
66 Gly Thr Ser Asn Ser Asp Thr Lys Val Gln Val Arg Ala Val Thr Ser
67          55          60          65          70
69 gcg tat ggc ctg tac tat acg cat gtg gat atc acg ttg aat acg atc 655
70 Ala Tyr Gly Leu Tyr Tyr Thr His Val Asp Ile Thr Leu Asn Thr Ile

```

## RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/963,521

TIME: 09:24:45

Input Set : A:\23282413.app

Output Set: N:\CRF3\12192001\I963521.raw

71		75		80		85	
73	acc atc ttc acc aac atc ggt gtg gag agg aag atg ccg gtc aac gtg	703					
74	Thr Ile Phe Thr Asn Ile Gly Val Glu Arg Lys Met Pro Val Asn Val						
75		90		95		100	
77	ttt cat gtt gtg ggc aag ttg gac acc aac ttc tcc aaa ctg tct gag	751					
78	Phe His Val Val Gly Lys Leu Asp Thr Asn Phe Ser Lys Leu Ser Glu						
79		105		110		115	
81	gtt gac cgt ttg atc cgt tcc att cag gct ggt gct acc ccg cct gag	799					
82	Val Asp Arg Leu Ile Arg Ser Ile Gln Ala Gly Ala Thr Pro Pro Glu						
83		120		125		130	
85	gtt gcc gag aaa att ctg gac gag ttg gag caa tcg cct gcg tct tat	847					
86	Val Ala Glu Lys Ile Leu Asp Glu Leu Glu Gln Ser Pro Ala Ser Tyr						
87	135		140		145		150
89	ggg ttc cct gtt gcg ttg ctt ggc tgg gca atg atg ggt ggc gct gtt	895					
90	Gly Phe Pro Val Ala Leu Leu Gly Trp Ala Met Met Gly Gly Ala Val						
91		155		160		165	
93	gct gtg ctg ttg ggt ggt gga tgg cag gtt tcc cta att gct ttt att	943					
94	Ala Val Leu Leu Gly Gly Gly Trp Gln Val Ser Leu Ile Ala Phe Ile						
95		170		175		180	
97	acc gcg ttc acg atc att gcc acg acg tca ttt ttg gga aag aag ggt	991					
98	Thr Ala Phe Thr Ile Ile Ala Thr Thr Ser Phe Leu Gly Lys Lys Gly						
99		185		190		195	
101	ttg cct act ttc ttc caa aat gtt gtt ggt ggt ttt att gcc acg ctg	1039					
102	Leu Pro Thr Phe Phe Gln Asn Val Val Gly Gly Phe Ile Ala Thr Leu						
103		200		205		210	
105	cct gca tcg att gct tat tct ttg gcg ttg caa ttt ggt ctt gag atc	1087					
106	Pro Ala Ser Ile Ala Tyr Ser Leu Ala Leu Gln Phe Gly Leu Glu Ile						
107	215		220		225		230
109	aaa ccg agc cag atc atc gca tct gga att gtt gtg ctg ttg gca ggt	1135					
110	Lys Pro Ser Gln Ile Ala Ser Gly Ile Val Val Leu Leu Ala Gly						
111		235		240		245	
113	ttg aca ctt gtg caa tct ctg cag gac ggc atc acg ggc gct ccg gtg	1183					
114	Leu Thr Leu Val Gln Ser Leu Gln Asp Gly Ile Thr Gly Ala Pro Val						
115		250		255		260	
117	aca gca agt gca cga ttt ttt gaa aca ctc ctg ttt acc ggc ggc att	1231					
118	Thr Ala Ser Ala Arg Phe Phe Glu Thr Leu Leu Phe Thr Gly Gly Ile						
119		265		270		275	
121	gtt gct ggc gtg ggt ttg ggc att cag ctt tct gaa atc ttg cat gtc	1279					
122	Val Ala Gly Val Gly Leu Gly Ile Gln Leu Ser Glu Ile Leu His Val						
123		280		285		290	
125	atg ttg cct gcc atg gag tcc gct gca gca cct aat tat tcg tct aca	1327					
126	Met Leu Pro Ala Met Glu Ser Ala Ala Ala Pro Asn Tyr Ser Ser Thr						
127	295		300		305		310
129	ttc gcc cgc att atc gct ggt ggc gtc acc gca gcg gcc ttc gca gtg	1375					
130	Phe Ala Arg Ile Ile Ala Gly Gly Val Thr Ala Ala Ala Phe Ala Val						
131		315		320		325	
134	ggg tgt tac gcg gag tgg tcc tcg gtg att att gcg ggg ctt act gcg	1423					
135	Gly Cys Tyr Ala Glu Trp Ser Ser Val Ile Ile Ala Gly Leu Thr Ala						
136		330		335		340	

## RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/963,521

TIME: 09:24:46

Input Set : A:\23282413.app

Output Set: N:\CRF3\12192001\I963521.raw

```

138 ctg atg ggt tct gcg ttt tat tac ctc ttc gtt gtt tat tta ggc ccc 1471
139 Leu Met Gly Ser Ala Phe Tyr Tyr Leu Phe Val Val Tyr Leu Gly Pro
140      345      350      355
142 gtc tct gcc gct gcg att gct gca aca gca gtt ggt ttc act ggt ggt 1519
143 Val Ser Ala Ala Ala Ile Ala Ala Thr Ala Val Gly Phe Thr Gly Gly
144      360      365      370
146 ttg ctt gcc cgt cga ttc ttg att cca ccg ttg att gtg gcg att gcc 1567
147 Leu Leu Ala Arg Arg Phe Leu Ile Pro Pro Leu Ile Val Ala Ile Ala
148 375      380      385      390
150 ggc atc aca cca atg ctt cca ggt cta gca att tac cgc gga atg tac 1615
151 Gly Ile Thr Pro Met Leu Pro Gly Leu Ala Ile Tyr Arg Gly Met Tyr
152      395      400      405
154 gcc acc ttg aat gat caa aca ctc atg ggt ttc acc aac att gcg gtt 1663
155 Ala Thr Leu Asn Asp Gln Thr Leu Met Gly Phe Thr Asn Ile Ala Val
156      410      415      420
158 gct tta gcc act gct tca tca ctt gcc gct ggc gtg gtt ttg ggt gag 1711
159 Ala Leu Ala Thr Ala Ser Ser Leu Ala Ala Gly Val Val Leu Gly Glu
160      425      430      435
162 tgg att gcc cgc agg cta cgt cgt cca cca cgc ttc aac cca tac cgt 1759
163 Trp Ile Ala Arg Arg Leu Arg Arg Pro Pro Arg Phe Asn Pro Tyr Arg
164      440      445      450
166 gca ttt acc aag gcg aat gag ttc tcc ttc cag gag gaa gct gag cag 1807
167 Ala Phe Thr Lys Ala Asn Glu Phe Ser Phe Gln Glu Glu Ala Glu Gln
168 455      460      465      470
170 aat cag cgc cgg cag aga aaa cgt cca aag act aat caa aga ttc ggt 1855
171 Asn Gln Arg Arg Gln Arg Lys Arg Pro Lys Thr Asn Gln Arg Phe Gly
172      475      480      485
174 aat aaa agg taaaaatcaa cctgcttagg cgtctttcgc ttaaatagcg 1904
175 Asn Lys Arg
177 tagaatatcg ggtcgatcgc ttttaaacac tcaggaggat ccttgccggc caaaatcacg 1964
179 gacactcgtc ccacccacaga atcccttcac gctggtgaag aggaaaccgc agccggtgcc 2024
181 cgcaggattg ttgccacctt ttctaaggac ttcttcgacg gcgtcacttt gatgtgcatg 2084
183 ctccggcgtt aacctcaggg cctgcgttac accaaggctg cttctgaaca cgaggaagct 2144
185 cagccaaaaga aggtacaaa gcggactcgt aaggcaccag ctaagaaggc tgctgctaag 2204
187 aaaacgacca agaagaccac taagaaaact actaaaaaga ccaccgcaaa gaagaccaca 2264
189 aagaagtctt aagccggatc ttatatggat gattccaata gctttgtagt tgttgctaac 2324
191 cgtctgccag tggatatgac tgtccacca gatggttagt atagcatctc cccagcccc 2384
193 ggtggccttg tcacggggct ttccccggt ctggaacaac atcgtggatg ttgggtcgga 2444
195 tggcctggaa ctgtagatgt tgcaccogaa ccatttcgaa cagatacggg tgttttgctg 2504
197 caccctgttg tctcactgc aagtgactat gaaggcttct acgagggtt ttcaaacgca 2564
199 acgctgtggc ctcttttcca cgatttgatt gttactccgg tgtacaacac cgattggtgg 2624
201 catgcgtttc gggaagtaaa cctcaagttc gctgaagcgg tgagccaagt ggcggcacac 2684
203 ggtgccactg tgtgggtgca ggactatcag ctggtgctgg ttcttgcat tttgcgccag 2744
205 atgcgccctg atttgaagat cggttttctt ctccacattc ccttcccttc cctgatctg 2804
207 ttccgtcagc tgc 2817
210 <210> SEQ ID NO: 2
211 <211> LENGTH: 489
212 <212> TYPE: PRT
213 <213> ORGANISM: Corynebacterium glutamicum

```

## RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/963,521

TIME: 09:24:46

Input Set : A:\23282413.app

Output Set: N:\CRF3\12192001\I963521.raw

215 &lt;400&gt; SEQUENCE: 2

```

216 Met Leu Ser Phe Ala Thr Leu Arg Gly Arg Ile Ser Thr Val Asp Ala
217   1           5           10           15
219 Ala Lys Ala Ala Pro Pro Pro Ser Pro Leu Ala Pro Ile Asp Leu Thr
220           20           25           30
222 Asp His Ser Gln Val Ala Gly Val Met Asn Leu Ala Ala Arg Ile Gly
223           35           40           45
225 Asp Ile Leu Leu Ser Ser Gly Thr Ser Asn Ser Asp Thr Lys Val Gln
226           50           55           60
228 Val Arg Ala Val Thr Ser Ala Tyr Gly Leu Tyr Tyr Thr His Val Asp
229   65           70           75           80
231 Ile Thr Leu Asn Thr Ile Thr Ile Phe Thr Asn Ile Gly Val Glu Arg
232           85           90           95
234 Lys Met Pro Val Asn Val Phe His Val Val Gly Lys Leu Asp Thr Asn
235           100          105          110
237 Phe Ser Lys Leu Ser Glu Val Asp Arg Leu Ile Arg Ser Ile Gln Ala
238           115          120          125
240 Gly Ala Thr Pro Pro Glu Val Ala Glu Lys Ile Leu Asp Glu Leu Glu
241           130          135          140
243 Gln Ser Pro Ala Ser Tyr Gly Phe Pro Val Ala Leu Leu Gly Trp Ala
244  145           150          155          160
246 Met Met Gly Gly Ala Val Ala Val Leu Leu Gly Gly Gly Trp Gln Val
247           165          170          175
249 Ser Leu Ile Ala Phe Ile Thr Ala Phe Thr Ile Ile Ala Thr Thr Ser
250           180          185          190
252 Phe Leu Gly Lys Lys Gly Leu Pro Thr Phe Phe Gln Asn Val Val Gly
253           195          200          205
255 Gly Phe Ile Ala Thr Leu Pro Ala Ser Ile Ala Tyr Ser Leu Ala Leu
256           210          215          220
258 Gln Phe Gly Leu Glu Ile Lys Pro Ser Gln Ile Ile Ala Ser Gly Ile
259  225           230          235          240
261 Val Val Leu Leu Ala Gly Leu Thr Leu Val Gln Ser Leu Gln Asp Gly
262           245          250          255
264 Ile Thr Gly Ala Pro Val Thr Ala Ser Ala Arg Phe Phe Glu Thr Leu
265           260          265          270
267 Leu Phe Thr Gly Gly Ile Val Ala Gly Val Gly Leu Gly Ile Gln Leu
268           275          280          285
270 Ser Glu Ile Leu His Val Met Leu Pro Ala Met Glu Ser Ala Ala Ala
271           290          295          300
273 Pro Asn Tyr Ser Ser Thr Phe Ala Arg Ile Ile Ala Gly Gly Val Thr
274  305           310          315          320
276 Ala Ala Ala Phe Ala Val Gly Cys Tyr Ala Glu Trp Ser Ser Val Ile
277           325          330          335
279 Ile Ala Gly Leu Thr Ala Leu Met Gly Ser Ala Phe Tyr Tyr Leu Phe
280           340          345          350
282 Val Val Tyr Leu Gly Pro Val Ser Ala Ala Ala Ile Ala Ala Thr Ala
283           355          360          365
285 Val Gly Phe Thr Gly Gly Leu Leu Ala Arg Arg Phe Leu Ile Pro Pro
286           370          375          380

```

## RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/963,521

TIME: 09:24:46

Input Set : A:\23282413.app

Output Set: N:\CRF3\12192001\I963521.raw

```

288 Leu Ile Val Ala Ile Ala Gly Ile Thr Pro Met Leu Pro Gly Leu Ala
289 385          390          395          400
291 Ile Tyr Arg Gly Met Tyr Ala Thr Leu Asn Asp Gln Thr Leu Met Gly
292          405          410          415
294 Phe Thr Asn Ile Ala Val Ala Leu Ala Thr Ala Ser Ser Leu Ala Ala
295          420          425          430
297 Gly Val Val Leu Gly Glu Trp Ile Ala Arg Arg Leu Arg Arg Pro Pro
298          435          440          445
300 Arg Phe Asn Pro Tyr Arg Ala Phe Thr Lys Ala Asn Glu Phe Ser Phe
301          450          455          460
303 Gln Glu Glu Ala Glu Gln Asn Gln Arg Arg Gln Arg Lys Arg Pro Lys
304 465          470          475          480
306 Thr Asn Gln Arg Phe Gly Asn Lys Arg
307          485
310 <210> SEQ ID NO: 3
311 <211> LENGTH: 1909
312 <212> TYPE: DNA
313 <213> ORGANISM: Corynebacterium glutamicum
315 <220> FEATURE:
316 <221> NAME/KEY: CDS
317 <222> LOCATION: (280)..(1746)
318 <223> OTHER INFORMATION: thrE-Gen
320 <400> SEQUENCE: 3
321 agcttgcatg cctgcaggtc gactctagag gatccccccc ctttgacctg gtgttattga 60
322 gctggagaag agacttgaac tctcaacctc cgcattacaa gtgcgttgcg ctgccaatg 120
323 cgccactcca gcaccgcaga tgctgatgat caacaactac gaatacgtat cttagcgtat 180
324 gtgtacatca caatggaatt cggggctaga gtatctggtg aaccgtgcat aaacgacctg 240
325 tgattggact ctttttcctt gcaaaatggt ttccagcgg atg ttg agt ttt gcg 294
326          Met Leu Ser Phe Ala
327          1          5
328 acc ctt cgt ggc cgc att tca aca gtt gac gct gca aaa gcc gca cct 342
329 Thr Leu Arg Gly Arg Ile Ser Thr Val Asp Ala Ala Lys Ala Ala Pro
330          10          15          20
331 ccg cca tcg cca cta gcc ccg att gat ctc act gac cat agt caa gtg 390
332 Pro Pro Ser Pro Leu Ala Pro Ile Asp Leu Thr Asp His Ser Gln Val
333          25          30          35
334 gcc ggt gtg atg aat ttg gct gcg aga att ggc gat att ttg ctt tct 438
335 Ala Gly Val Met Asn Leu Ala Ala Arg Ile Gly Asp Ile Leu Leu Ser
336          40          45          50
337 tca ggt acg tca aat agt gac acc aag gta caa gtt cga gca gtg acc 486
338 Ser Gly Thr Ser Asn Ser Asp Thr Lys Val Gln Val Arg Ala Val Thr
339          55          60          65
340 tct gcg tac ggt ttg tac tac acg cac gtg gat atc acg ttg aat acg 534
341 Ser Ala Tyr Gly Leu Tyr Tyr Thr His Val Asp Ile Thr Leu Asn Thr
342          70          75          80          85
343 atc acc atc ttc acc aac atc ggt gtg gag agg aag atg ccg gtc aac 582
344 Ile Thr Ile Phe Thr Asn Ile Gly Val Glu Arg Lys Met Pro Val Asn
345          90          95          100
346 gtg ttt cat gtt gta ggc aag ttg gac acc aac ttc tcc aaa ctg tct 630

```

**VERIFICATION SUMMARY**

DATE: 12/20/2001

PATENT APPLICATION: US/09/963,521

TIME: 09:24:47

Input Set : A:\23282413.app

Output Set: N:\CRF3\12192001\I963521.raw